Impaired Water Identification Rule

Stakeholder Meeting to discuss potential revisions

June 22, 2004

Meeting Agenda

Goal: Present issues with the IWIR, take comments, consider options to improve rule

- Summary of 303(d) Listing Process and Issues with the Current Rule
  Susan Craig
- Comparison and Discussion of Current Methods and Alternatives: What do we want to improve?
  - Binomial Approach
    Melanie Diroll
  - Chronic Aquatic and Wildlife Standards
    Diana Marsh
  - *Escherichia coli*
    Melanie Diroll
  - Planning List and Narratives (time permitting)
    Melanie Diroll
- Final Comments and Discussion
  All

The Integrated 305(b) Assessment and 303(d) Listing Report – Federal Requirements

- **Section 305(b) of the Clean Water Act**
  Assessment of all surface waters every two years (even-numbered years)

- **Section 303(d) of the Clean Water Act**
  List of impaired waters every two years

- **Submitted together as an Integrated Report**
  - Arizona submits report to EPA Region IX (San Francisco, CA) for approval of the 303(d) List
  - EPA can approve, disapprove, or partially disapprove list
  - EPA can add or remove surface waters and/or pollutants

Arizona's Integrated Report – State Requirements

- Statutory Requirements
  - Arizona Revised Statutes (A.R.S.) § 49-232
  - Required adoption of rule for identifying impaired surface waters
- Impaired Water Identification Rule
  - Arizona Administrative Code (A.A.C.) R18-11, Article 6
  - Methods for determining impairment
  - Credible data requirements
- The rest of the report
  - What about waters where we don't have exceedances?
  - Methods for determining attainment established internally, using EPA guidance, and made available for public review (in draft Integrated Report)

How does ADEQ prepare the Integrated Report?

- Consider last five years of data

  2004 report – 1/1/98 through 12/31/02
  2006 report – 1/1/00 through 12/31/04

- Assess data using standards established in Arizona Surface Water Quality Standards

  Aquatic and Wildlife, Fish Consumption, Full/Partial Body Contact, Agricultural Livestock Watering, Agricultural Irrigation

- Designated Uses can be assessed as:

  Attaining, Inconclusive, Impaired, Not Attaining

- Stream Reaches and Lakes can be assessed as:

  Attaining all uses, Attaining some uses, Inconclusive, Impaired, Not attaining

What has happened since the IWIR was adopted in 2002?

- 2002 303(d) List completed using IWIR methods

  EPA Region IX partially disapproved and added 19 surface waters/pollutants to the list

  (letter to Karen Smith dated (received) March 7, 2003)

- Draft 2004 List completed

  EPA Region IX included a preliminary list of waters that will be added in its comment letter

  (letter to Linda Taunt dated Dec. 8, 2003)

What are EPA's Issues?

- The Binomial Approach

  Listing based on 10% exceedance rate at a 90% confidence level may overestimate error, not consistent with WQ standards

- Minimum number of exceedances and minimum sample size

  5 exceedances, 20 total samples

- Narrative Standards

  We must adopt implementation procedures before we can make listings for violations such as fish consumption advisories, excessive bottom deposits.

What are ADEQ's Issues?

- Chronic Aquatic and Wildlife Standards
  - Listing based on more than one exceedance in the 5-year assessment period
  - Similar to acute listings (more than one exceedance in the last 3 years)
- *Escherichia coli*
  - Analytical methods provide an estimate of bacterial density
  - Listing based on more than one exceedance of the single sample max in the last 3 years
  - 30-day timeframe for the geometric mean no longer specified in standards
- Planning List (time permitting)
  - Planning List has several purposes
  - Caused some confusion in 2004

Binomial Approach

Current IWIR Method

- Statistical method for determining impairment
  10% exceedance rate at a 90% confidence level using equation (R18-11-605(1)(b))
- Applies to conventionals (D.O., pH) and most standards for human health and agricultural uses
- Excludes chronic and acute A&W, nitrate, *E. coli*, and statistically-based standards (such as nutrient annual mean)
- Establishes minimum sample size of 20 to make a listing, with a minimum of 5 exceedances, collected over three or more sampling events

EPA Reasons for Disagreement

- Our application of binomial approach overestimates error
- 10% exceedance rate accounts for sampling error
- Two to three exceedances sufficient to make a listing for smaller sample sizes
- Not consistent with how WQ standards are expressed

Listing Error

Type I Error

Listing a water that is <u>not impaired</u>

Type II Error

Not listing a water that <u>is impaired</u>

Type I Error:
Listing waters as impaired that are not, in fact, impaired

- 10% exceedance rate is recommended in CALM for conventionals
- The 10% is not an "allowable" exceedance rate, but is an assessment method that accounts for sampling error, uncertainty and manages Type I error. Adding binomial on top of a 10% exceedance rate allows a much greater % exceedances.
- EPA: "Because the binomial approach already accounts for and directly manages uncertainty associated with assessments based on small sample sizes... it would be inappropriate to apply the 10% exceedance rate directly within the context of a binomial assessment approach. To use a 10% test in a binomial assessment context would, in essence, result in 'double counting' of allowances intended to limit Type I error."
  (letter to Shirley Conard, dated Oct. 4, 2001)

Type II Error:
Not listing waters that are, in fact, impaired

- CALM: "Smaller sample sizes are prone to yield erroneous attainment decisions because they have a low probability of detecting WQS exceedances unless they are large and pervasive."

- Type II error managed through larger sample sizes. If a water quality exceedance truly exists, the probability of detecting it in a small sample size is very low. In other words, if you find a problem in a small sample set (i.e. 3 of 3 exceedances), chances are very good that a serious water quality problem does indeed exist.

- Therefore, EPA advises listing when sufficient exceedances have occurred, regardless of total samples taken. Recommend two or three instead of five.

Binomial compared with 10% Method

# samples	# exceeds for 303(d) listing
1-2	---
3-19	--- (EPA lists)
20–29	5-6
30–39	6-7
40-49	7-8

Binomial

# samples	# exceeds for 303(d) listing
1-2	---
3-19	3
20-29	3
30-39	4
40-49	5

10% method

Other Options?

- Consider magnitude of exceedance
  - Would we list at <10% if magnitude high?
  - What would we define as “high?”
- Modify the binomial
  - Change confidence level or exceedance rate?

Assessment of Chronic Aquatic and Wildlife Standards

Current IWIR Method

- Chronic standards
  - Protect aquatic life and wildlife from long-term effects of low level exposure to toxic pollutants. Criteria are based on a 4-day exposure.
  - IWIR R18-11-605.C.3.b.iii – Listing based on > 1 exceedance of aquatic and wildlife chronic criteria. (Using a 5-year assessment period.)

- Acute standards
  - Protect from short-term effects of high concentrations of toxic pollutants. Criteria are based on a 1-hour exposure.
  - 2004 303(d) listing was based on >1 exceedance in a 3-year period.

Assessment of Chronic Aquatic and Wildlife Standards

Possible Alternatives

- Listing if >1 exceedances in a 3-year period.
  - EPA Technical Support Document (1991), CALM guidance, and EPA Assessment Guidance (2004) supports this.
- Listing if >1 exceedances in a 5-year period.
  - 2004 method
- Listing if >10% exceedance rate.
  - Same method as proposed to replace binomial.
- Listing if median (mean or central tendency) exceeds standards.
  - Several technical problems using this
- Chronic standard screening value is calculated (e.g., chronic standard is multiplied by 1.5) and compared to results. Listing if >1 exceedance in a 3-year period.

Example: Assessing Dissolved Copper

Date	Result (µg/L)	Hardness (mg/L)	Calculated Standard	Assessment
6/25/01	31	400	29.28	Exceeds
7/1/01	29	200	16.19	Exceeds
8/4/01	31	400	29.28	Exceeds
7/1/04	19	250	19.59	
8/12/04	25	312	23.68	Exceeds
9/15/04	25	380	28.02	

If apply >1 exceedance in 3-years, it would not be listed -- only 1 exceedance in the last 3 years.

Example: Assessing Dissolved Copper

Date	Result (µg/L)	Hardness (mg/L)	Calculated Standard	Assessment
6/25/01	31	400	29.28	Exceeds
7/1/01	29	200	16.19	Exceeds
8/4/01	31	400	29.28	Exceeds
7/1/04	19	250	19.59	
8/12/04	25	312	23.68	Exceeds
9/15/04	25	380	28.02	

If apply >1 exceedance in 5-years, then it meets the listing criteria -- 4 exceedances occurred.

Example: Assessing Dissolved Copper

Date	Result (µg/L)	Hardness (mg/L)	Calculated Standard	Assessment
6/25/01	31	400	29.28	Unknown
7/1/01	29	200	16.19	Unknown
8/4/01	31	400	29.28	Unknown
7/1/04	19	250	19.59	Unknown
8/12/04	25	312	23.68	Unknown
9/15/04	25	380	28.02	Unknown

Cannot use the mean of the results because cannot determine what hardness would apply, and therefore, what standard would apply.

Example: Assessing Dissolved Copper

Date	Result (µg/L)	Hardness (mg/L)	Calculated Standard	Assessment
6/25/01	31	400	29.28	Unknown
7/1/01	29	200	16.19	Unknown
8/4/01	31	400	29.28	Unknown
7/1/04	19	250	19.59	Unknown
8/12/04	25	312	23.68	Unknown
9/15/04	25	380	28.02	Unknown

If apply median, the median result is between 25 and 29. But what hardness would apply?

Example: Assessing Dissolved Copper

Date	Result (µg/L)	Hardness (mg/L)	Calculated Standard	Assessment
6/25/01	**31**	400	29.28	Exceeds
7/1/01	29	200	16.19	Exceeds
8/4/01	**31**	400	29.28	Exceeds
7/1/04	19	250	19.59	
8/12/04	**25**	312	23.68	Exceeds
9/15/04	**25**	380	28.02	

If apply mode, there are 2 modes, 31 and 25. Three of the four results exceeded standards, should we do a listing?

Example: Assessing Dissolved Copper

Date	Result (µg/L)	Hardness (mg/L)	Calculated Standard	Screen Value 1.5	Assessment
6/25/01	31	400	29.28	43.92	
7/1/01	29	200	16.19	24.28	Exceeds
8/4/01	31	400	29.28	43.92	
7/1/04	19	250	19.59	29.38	
8/12/04	25	312	23.68	35.52	
9/15/04	25	380	28.02	42.03	

If apply screening value technique, only 1 exceedance, and none in last 3 years. Would not list. How defend the 1.5?

Chronic Screening Values

- Montana's chronic listing criteria:
  - If >1 acute standard exceedance in 3 years (don't bother looking at chronic exceedances).
  - If chronic screening value (standard times 2.5) was exceeded >1 in 5 years.
  - If less than 10 samples and chronic screening value (standard times 1.1) was exceeded >1 in 5 years.

Chronic Screening Values

- New Mexico's chronic listing criteria is:
  - If chronic screening value (standard times 1.5) was exceeded >1 in assessment period (5 years).

- Utah's chronic listing criteria is:
  - If < 10 samples and chronic screening value (standard times 1.75) was exceeded >1 in a 3-year period.
  - If > 10 samples and >2 exceedances in a 3-year period.

Chronic Screening Values

Different EPA regional offices are reviewing these. Montana's has been accepted by Region VIII. Not known whether and under what conditions Region IX might accept this concept.

Region VIII indicated their support of screening values is based on:

- Chronic standards based on a 4-day average of data; however, grab samples (single samples) are used to represent this period.

- Lack of EPA guidance concerning how to deal with small data sets (less than 10 samples).

What would the multiplication factor be? What is the technical support for this factor?

Example: Assessing Selenium

Chronic standard = 2 µg/L

4 of 80 samples exceeded

2 in the last 3 years

Results = 7.3, 2.8

>1 exceedance in 3 years.	List – 2 in 3 years
>1 exceedance in 5 years.	List – 4 in 5 years
> 10% of samples exceed.	Not list – only 5%
Mean of sample results exceeds standard.	Not list – only 5%
Magnitude – screening value of 1.5 times standard and > 1 exceedance in 3 years.	Not list – 1 in 3 years

Working with larger number of samples and standards that are not hardness dependent.

Escherichia coli: Analytical Accuracy

- Analytical methods provide an estimate of bacterial density
- Full Body Contact = 235 CFU/100ml (single sample max)
  - Colilert (Quanti-tray method)

    Result of 240

    Actual result between 146 and “infinite” (95% confidence level)
  - Fermentation tube method

    Result of 240

    Actual result between 100 and 940 (95% confidence level)
- How to consider for listing?

  For 2004, did not use a result <300 to make a listing

Escherichia coli: Listing based on single sample max

FBC standard = 235 CFU/100 ml

PBC standard = 576 CFU/100 ml

- Current IWIR method – listing based on more than one exceedance in the last 3 years of available data (R18-11-605(D)(2)(a))
- Appropriate for large and small sample sizes, and for large and small waterbodies?

Examples: Small and Large Sample Sets

Current method

ADEQ Ambient Monitoring

- 4 quarters
- 2 of 4 exceedances
- Listed -- Reasonable

10% method

- Would need 3 exceedances – Not listed
- Not reasonable for small sample sets
- If decrease minimum exceedance to 2?

Current method

Swimming areas

- 500 samples
- 2 of 500 exceedances
- Listed – Reasonable?

10% method

- Would not be listed
- Reasonable option?

Large Sample Sets: Another example

Current method

- Swimming area
- 500 samples, 45 of 500 exceed
- Listing is reasonable

10% method

- 9% exceed
- Would not be listed
- Reasonable option? – No
- Could establish a maximum value or use a smaller percentage

Example: Large Lakes

- 3 out of 200 exceedances, at three separate beaches several miles apart. Should we make a listing?
- In 2004, considered beaches separately, assessed as "inconclusive." Do we want to establish in rule? What distance should be considered "separate?"

Summary: *E. coli* single sample max

- Frequency of monitoring
  - Consider different listing methods for different sample sizes?
- Size of the waterbody
  - Consider different listing methods for waterbodies of variable sizes?
- Or, can we make one method work?

Escherichia coli: Assessing geometric mean standard

Geo mean standard = 126 CFU/100ml

- IWIR lists when >1 exceedance of a 30-day geometric mean occurs – based on old standards (R18-11-605(D)(2)(b))

- New standards (2002) require a minimum of 4 samples for the *E. coli* geometric mean (125 CFU/100ml for FBC)
  - Should the geo mean consider all samples, during swimming and non-swimming seasons? Doesn't work if we need more than one exceedance as required in IWIR.
  - How do we make it work? Should we keep the 30-day in IWIR for swimming areas where there is sufficient data? If less data, can we also consider the entire swimming season?

Planning List

- Currently required in IWIR (R18-604(D)(2)) for several reasons, including:
  - Exceedance occurred, not enough data to assess impaired
  - Evidence of a narrative violation
  - Data does not meet credible data requirements
  - TMDL follow-up monitoring
  - Insufficient sample results over the lab detection limit
  - Exceedances due to pollution but not a pollutant
- ADEQ also places waters on Planning List that are lacking sufficient data to make an assessment.
- Do we need it in rule at all? We will still track all waters in the five categories.
- Leave in rule but add clarification?

Final Discussion and Comments: What would you like to improve?

Submit written comments	July 9, 2004
Meetings to discuss options	July - September 2004
Draft rule for review	September 2004
Publish rule for review	September 2004
Meetings to discuss rule	October 2004
Publish notice of proposed rulemaking	November 2004
Oral Proceedings	January 2005
File Notice with GRCC	February 2005
File Rule with Sec. of State	March 2005

Next Meeting: end of July, beginning of August

Comments to:
Melanie Diroll, Assessment Coordinator
md1@ev.state.az.us
(602) 771-4616

Arizona Department of Environmental Quality
1110 W. Washington Street
Hydrologic Support and Assessment, 5415A-1
Phoenix, AZ 85007